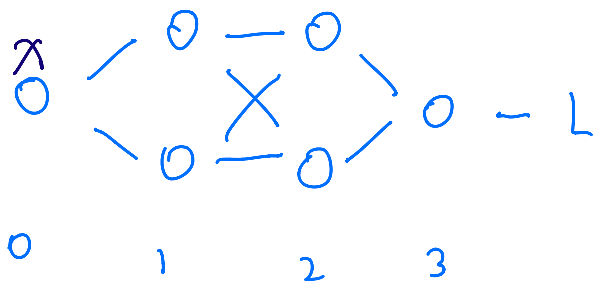

$$Z^{i} = (W^{i})^{T} a^{[i-1]} + b$$

$$L(a,y) = -y\log a - (1-y)\log(1-a)$$

$$J(a,y) = \frac{1}{m}\sum_{i=1}^{m} L(a,y)$$

x, $W^{[1]}$, $b^{[1]}$ $\rightarrow$ $Z^{[1]} = W^{[1]}x + b^{[1]} \rightarrow a^{[1]} = \sigma(Z^{[1]})$

$\rightarrow Z^{[2]} = W^{[2]}a^{[1]} + b^{[2]} \rightarrow a^{[2]} = \sigma(Z^{[2]})$

$\rightarrow Z^{[3]} = W^{[3]}a^{[2]} + b^{[3]} \rightarrow a^{[3]} = \sigma(Z^{[3]})$

$$dZ^{[3]} = \frac{\partial L}{\partial a^{[3]}}\frac{\partial a^{[3]}}{\partial Z^{[3]}} = \left(-\frac{y}{a^{[3]}} + (1-y)\frac{1}{1-a^{[3]}}\right)a^{[3]}(1-a^{[3]}) = a^{[3]} - y$$

$$dW^{[3]} = dZ^{[3]} \cdot \frac{\partial Z^{[3]}}{\partial W^{[3]}} = dZ^{[3]} \cdot (a^{[2]})^{T}$$

$\frac{\partial x^{T}A}{\partial x} = A^{T}$ （分子布局）

$$db^{[3]} = dZ^{[3]}\frac{\partial Z^{[3]}}{\partial b^{[3]}} = dZ^{[3]}$$

$$dZ^{[2]} = dZ^{[3]} \cdot \frac{\partial Z^{[3]}}{\partial a^{[2]}} \cdot \frac{\partial a^{[2]}}{\partial Z^{[2]}} = (W^{[3]})^{T} dZ^{[3]} * g^{[2]\prime}(Z^{[2]})$$

$(n_2, n_3) \cdot (n_3, 1) * (n_2, 1)$

$$dW^{[2]} = dZ^{[2]} \cdot (a^{[1]})^{T}$$

$$db^{[2]} = dZ^{[2]}$$

$$dZ^{[1]} = dZ^{[2]} \cdot \frac{\partial Z^{[2]}}{\partial a^{[1]}} \cdot \frac{\partial a^{[1]}}{\partial Z^{[1]}}$$

$$= (W^{[2]})^T \cdot dZ^{[2]} * g^{[1]'}(Z^{[1]})$$

$$(n_1, n_2) \cdot (n_2, 1) * (n_1, 1)$$

$$dW^{[1]} = dZ^{[1]} \cdot X^T$$

$$db^{[1]} = dZ^{[1]}$$